Semantic Prosody and its Effect on EFL learners’ Lexical Cohesion

Mohsen Jafarpour (m_jafarpour@ut.ac.ir)
Tehran University, Iran

Majid Nemati (nematim@ut.ac.ir)
Tehran University, Iran

Aliakbar Khamijani farahani
Tehran University, Iran

Abstract
The first part of the study focused on the notion of semantic prosody (SP). It was demonstrated that SP should become an integral part of dictionary definitions to distinguish near-synonyms. The second part of the study examined presenting SP of near-synonyms and its effect on learners’ lexical cohesion in an English as a foreign language (EFL) context. To achieve this, three male intact classes at advanced level were selected. Two of these classes were randomly selected as experimental groups (A & B) and another one as a control group. Experimental group A was provided with regular teaching of SP of near-synonymous lexical items through the data-driven learning (DDL) approach. Experimental group B was only informed of SP of the same lexical items. However, the control group was taught traditionally without being exposed to the concept of SP. In order to measure lexical cohesion, pretests and posttests of task 2 of IELTS writing examination were administered before and immediately after the study. One-way ANOVA and Post-hoc Tukey Test were used to compare means of test scores between groups. The significant effect of semantic prosody on lexical cohesion was confirmed. Finally, implications on integrating SP into language pedagogy were discussed.

Keywords: data-driven learning, near-synonyms, semantic prosody, lexical cohesion

Introduction
Nowadays, vocabulary is a priority in language instruction (Richards & Rodgers, 2001) and learning it is “the first step to learn a foreign language” (Alhamami, 2016, p. 87). Moreover, its importance has been acknowledged in many second language (L2) studies (e.g. August, Carlo, Dressler, & Snow, 2005; Schmitt, Jiang, & Grabe, 2011). However, an area where L2 learners commonly exhibit a number of errors is lexical use (Watter, 1992). Since over 60% of English vocabulary consist of synonyms (henceforth referred to near-synonyms), it is important for English learners to get help to be able to discriminate them (Watter, 1992; Zhao, 2017).
Although words in a near-synonym cluster have close meaning, they are not necessarily interchangeable in practical use because their specific usages and collocational constraints are different (Yu, Shih, Lai, Yeh, & Wu, 2010). In fact, “the concepts that are common to all the near-synonyms in a cluster could be part of their main meaning, while those that associate only with one near-synonym could be part of their implied nuances of meaning” (Inkpen, 2007, p. 15). L2 teachers are warned that lack of semantic appropriateness, due to inappropriate lexical choice among near-synonyms, leads to unclear communication and damaging social consequences (Lee & Liu, 2009).

To choose among near-synonyms, language learners rely on dictionaries and thesauri “without being aware of the subtle implications embedded in contexts” (Lee & Liu, 2009, p. 206). Since these implications are “largely uncaptured by dictionary definitions” (Guo et al., 2011, p. 417), learners remain negligent of semantic distinctions among near-synonyms (Lee & Liu, 2009). Therefore, near-synonyms need to be devoted special attention.

Closely linked to the near-synonym concept is the term SP, which develops a new method to differentiate between near-synonyms (Fan, 2010; Hunston, 2007; Lee & Liu, 2009) and yields insight into teaching them (Zhang, 2010).

### Semantic Prosody

As a concept arisen out of corpus linguistics, SP has been studied for at least two decades. It is “deeply tied to the phenomenon of collocational sequence of lexical items” (Elahi & Rahbar, 2018, p. 75). It refers to a sort of connotative meaning (positive, neutral, or negative) that a word takes due to its consistent collocations. It allows us to understand words and their meaning along with their attitudinal nuances of use (Sorli, 2013). Words with semantic prosody “color ambiguous concepts with evaluative meaning” (Hauser & Schwars, 2018, p. 12). For example, the verb CAUSE is almost always associated with words such as difficulty, war, death, problem, and damage which are unpleasant and negative. Consequently, CAUSE has a negative SP. On the other hand, the verb BRING ABOUT, near-synonym of the verb CAUSE, is usually followed by positive concepts such as improvement and significant. Therefore, BRING ABOUT has a positive SP (Xiao & McEnery, 2006).

Table 1 (adapted from Xiao & McEnery, 2006) gives some examples of lexicons whose conditions have already been determined by different linguists.
Table 1
Examples of semantic prosodies

<table>
<thead>
<tr>
<th>Author</th>
<th>Negative prosody</th>
<th>Positive prosody</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sinclair (1991)</td>
<td>BREAK out</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HAPPEN</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SET in</td>
<td></td>
</tr>
<tr>
<td>Louw (1993, 2000)</td>
<td>[be] bent on build up of [intransitive]</td>
<td>BUILD up a [transitive]</td>
</tr>
<tr>
<td></td>
<td>END up verbing</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GET oneself verbed a recipe for</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CAUSE</td>
<td>Career</td>
</tr>
<tr>
<td></td>
<td>FAN the flame signs of underage teenagers (s)</td>
<td></td>
</tr>
<tr>
<td>Partington (1998)</td>
<td>COMMIT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PEDDLE/peddler</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dealings</td>
<td></td>
</tr>
<tr>
<td>Hunston (2002)</td>
<td>SIT through bordering on</td>
<td></td>
</tr>
</tbody>
</table>

Although lexical items may be associated with the same type of SP, their prosodic strengths vary considerably (Wei & Li, 2014). Moreover, in a SP, a word by itself is not explicitly positive or negative, rather it is its collocates that have a positive or negative semantic association (Zhang, 2010). Therefore, SP is not recognizable from words alone, but requires those words to be used by a particular set of participants to acquire a particular prosodic effect (Philip, 2010).

Violation of a SP leads to “expressions of insincerity as well as of irony” (Louw & Milojkovic, 2016, p. 54). In fact, if deliberately “there is sufficient distance between the expected collocation and the combination of words proposed by the author, the result is irony”. However, when the speaker is not aware of irony, the irony is unintentional and the real attitude of the speaker is revealed unconsciously (Louw & Chateau, 2010, p. 757).

**Semantic Prosody Instruction**

Researchers (Xiao & McEnery 2006; Zhang, 2009 among others) have observed that choosing inappropriate words among near-synonyms due to lack of knowledge of SP is very common among EFL learners. Neither intuition nor introspection is practical and trustworthy to inform them of SP of lexical items (Sardinha, 2000; Stewart, 2010). To EFL learners it is not feasible to find out the prosodic behavior of a word through “a priori intuition” (Louw & Chateau, 2010, p. 756). SP can be revealed only by corpus analysis (Lee & Liu 2009; Louw, 2000; Louw & Chateau, 2010) “rather than intuitively” (Louw, 2000, p. 3). That is, it “does not belong to speakers’ conscious knowledge of a language” (Zhang, 2009, p. 3).
Nonetheless, native speakers possess the knowledge of semantic prosody subconsciously (Zhao, 2017). This makes them to be able to understand the effects of it without being able to explain these effects (Louw & Chateau, 2010). Moreover, when SP is violated, “native speaker intuition certainly can detect the usage of a word at odds with its semantic prosody” (Xiao & McEnery, 2006, p. 126). That is, “semantic prosodies are part of all [native] readers’ prior knowledge” (Louw, 2000, p. 3).

However, non-native speakers might observe SP only by interpreting large numbers of instances of a word usually through corpora analysis (Zhang, 2009). In fact, regarding ignorance of SP, it is the EFL learner who represents inappropriate vocabulary choice. That is, “learners’ L2 intuition...is inevitably less reliable than their L1 intuition” (Xiao & McEnery, 2006, p. 126). Hence, “a data-based research of semantic prosody is needed” (Zhao, 2017, p. 436). The best tool for learning prosodic behavior of vocabulary is keyword-in-centre (KWIC) concordances “as these allow the learner to observe repeated patterns and meanings, and thus help them to become aware of collocation and semantic prosody” (Xiao & McEnery, 2006, p. 126). Furthermore, corpus study can give EFL learners the opportunity to surpass the native speakers’ intuition in judgment on SP (Louw & Chateau, 2010). The pedagogical approach that studies large amounts of linguistic data (corpora) by software programs named concordancers, usually on computers, in order to identify regular patterns, is called DDL. In effect, it is through DDL that patterns of SP can be discovered (Reinhardt, 2010).

Working out words from the context to decide which words go with which entities and observe the collocations that accompany a specific word appears to be a revolutionary way to discover SP of words. In fact, SP is often hidden from human intuition and so can only be explored by a DDL approach (Zhang, 2010).

**Data-Driven Learning and Semantic Prosody**

As a concept arisen from corpus linguistics, DDL “focuses on the typical behaviour of individual lexical items as observed using ‘key word in context’ concordance lines” (Hunston, 2007, p. 249). It isolates common patterns in authentic language samples with software programs called concordancers (Hadley, n.d.).

To Boulton (2011), DDL activities use a constructive approach “to inductive, problem-solving discovery learning from naturalistic pattern-recognition in authentic language data” (p. 2). In DDL, teachers should be a director and coordinator. Students are encouraged to explore linguistic texts. It helps them find patterns, rules, and consequently generalizations in linguistic context (Johns, 1991).

To illustrate the way DDL works, Figure 1 shows all 14 selections of instances of the verb RESORT TO in Brown + BNC Written Corpora with the keyword in center. With regard to the words indicating what entity is “resorted to”, in most cases, this entity is one which would normally be considered to be unfavorable: secrecy, taking hostages, military means,
hysterical repressions, and so on. Therefore, it can be concluded that the word RESORT TO has a negative SP.

Figure 1. All 14 selections of instances of the verb RESORT TO in Brown + BNC Written Corpora

Bednarek (2008) warns that:

(1) SP of a word in different types of texts will be different and it “is probably context-, genre- and domain dependent” (p. 123).

(2) Labeling an item as having a positive/ negative SP is a subjective matter. That is, the percentage of collocates identified as positive/ negative may be different from researcher to researcher.

(3) Lexical items with different word classes may also have distinct semantic prosodies. Thus, a verb may have different SP than its noun form.

As mentioned above, traditionally, teachers teach vocabulary by providing near-synonyms, among other things, and in doing so they fail to point out the distinguishing features among them. Similarly, in dictionaries words are usually defined by providing near-synonyms
without giving any information about SP. As a result, learners fail to identify the distinguishing features of near-synonyms, and use words inappropriately (Watter, 1992). Wrong choice of near-synonyms conveys undesired connotations, implications, and attitudes (Inkpen, 2007; Inkpen & Hirst, 2006). Therefore, the traditional practice of vocabulary by offering near-synonyms should be used with caution.

Xiao and McEnery (2006) demonstrated that near-synonyms are normally not interchangeable and “teachers, learners, and lexicographers have been advised not to use words with close meanings (near synonyms) at the expense of focusing on connotative meanings (semantic prosodies)” (Ahmadian, Yazdani, & Darabi, 2011, p. 288). Furthermore, since advanced learners have serious problems with near-synonyms (Hemchua & Schmitt, 2006), the knowledge of SP for EFL vocabulary instruction is essential (Inkpen & Hirst, 2006; Zhang, 2009). However, this knowledge is neglected by L2 learners (Ahmadian et al., 2011). Consequently, misuse of lexical items, particularly the choices among near synonyms, calls for more attention and treatment in L2 lexical learning (Lee & Liu, 2009; Inkpen, 2007).

Additionally, Halliday and Hasan (1976) argue that near-synonyms play a key role in lexical cohesion. To them, cohesion is a semantic concept, which refers to “relations of meaning that exist within the text, and define it as a text” (p. 4). It is assumed “as a powerful tool in discourse production and interpretation” (Tanskanen, 2006, P. 27).

Cohesion occurs where the interpretation of some element in the discourse is dependent on that of another. It is expressed through grammar and vocabulary. The cohesive effect that is achieved through the selection of vocabulary is called lexical cohesion. Moreover, there is a consensus among researchers that lexical cohesion consists of two different types of relation. (Tanskanen, 2006). These two main subclasses of lexical cohesion, as identified by Halliday and Hassan (1976), are reiteration and collocation.

Halliday and Hasan (1976) contend that reiteration involves the repetition of a lexical item, the use of near-synonyms, or super-ordinates, and a general word to refer back to a lexical item. Moreover, “any two lexical items having similar patterns of collocation … will generate a cohesive force if they occur in adjacent sentences” (p. 286). Therefore, it seems that SP performs a role in creating lexical cohesion. The concept of SP is relevant to cohesion since it is related to both collocation and reiteration. It is related to collocation because SP of a word is the result of its consistent collocation. It is also partly related to reiteration since SP is the main distinguishing feature of near-synonyms.

It should be noted that since cohesive devices are on the surface of the text, they can be observed, counted and analyzed and are therefore objective (Tanskanen, 2006, P.21). This makes the assessment of lexical cohesion easier.

However, little research has been carried out about SP (Zhang, 2010). Moreover, no study has been reported to investigate whether introducing SP has significant effect on the improvement of lexical cohesion. Therefore, this study addresses following null hypotheses:
(1) Teaching SP of lexical items through the DDL approach does not have any statistically significant effect on Iranian learners’ lexical cohesion in an EFL context.
(2) Knowledge of semantic prosody of lexical items does not have any statistically significant effect on Iranian learners’ lexical cohesion in an EFL context.
(3) There is no significant difference between the effect of teaching SP of lexical items through the DDL approach and the effect of knowledge of SP of lexical items on Iranian learners’ lexical cohesion in an EFL context.

Materials and Methods

This study followed a quasi-experimental and pretest-posttest design. Due to some organizational issues, intact classes were selected for carrying out the present study.

Participants

To carry out this study, 52 advanced participants were assigned as one control group (N=18), and two experimental groups: group A (N=17) and group B (N=17). All participants were chosen from Novin English Institution in Talesh, Iran, and were native speakers of Persian. In fact, three intact classes were selected. In order to make sure that the participants were homogeneous, the Oxford Placement Test (2004) was administered to the participants and those learners whose scores deviated one standard deviation below the mean on the test were excluded. Two of three classes were randomly selected as experimental groups (A & B) and the other class was selected as a control group.

Participants were male students. They had already studied English for 7 to 9 years, with a mean of 8 years. They were studying American English File 5, 2nd edition (Latham-Koenig & Oxenden, 2013). The main reason for choosing these learners, i.e., advanced learners, was that they had a greater chance to improve their lexical cohesion. The classes were held two sessions per week. Each session took 90 minutes out of which 20 minutes was devoted to working on vocabulary. This study was conducted for 6 months.

Data Collection Instruments

Two IELTS writing tasks were administered before and after the study as pretests and posttests to compare the participants’ lexical cohesion. The students were given a topic to write about during 40 minutes. The students were supposed to present their point of view with convincing evidence, write in a style that is easy to follow and cohesive, and use English accurately and appropriately.

Oxford Placement Test (2004) was administered to both the experimental groups and the control group before the study to choose homogenous participants. Moreover, American English File 5, 2nd edition (Latham-Koenig & Oxenden, 2013) was used in this study as the course book.
To find semantic behavior of near-synonyms, the Brown University Corpus of American English (Francis & Kucera, 1964) and British National Corpus (BNC) were chosen. They were chosen because they were easily available, they had been used in previous studies, and (as it was the first time that the students were supposed to use a corpus for finding out the SP of lexical items) it was manageable and easy enough to use.

Procedures

In experimental group A, the participants were provided with regular teaching of SP of lexical items through the DDL approach. For words (mainly verbs) that students encountered in their course book, they were asked to find its near-synonyms in a thesaurus. However, not all words and their near-synonyms were studied. The teacher helped them select active one. In the next step, they were asked to find SP of near-synonyms for comparative purposes, so that they knew, in a synonym set, which word had a positive, which word had a neutral and which word had a negative SP. They were also taught how to use corpora and concordancer in order to find SP of near-synonyms by themselves.

The positive, neutral, and negative prosodies were taken like Xiao and Mcenery (2006), corresponding to Partington’s (2004) favorable, neutral, and unfavorable prosodies. That is, in practical terms, the phenomenon of SP was classified into three major types: a positive one (+SP), a negative one (-SP) and a neutral one (~SP).

In order to provide a favorable environment for adopting the DDL approach, the institution was equipped with wireless internet connection. The participants had access to computers, the online corpora, and their concordancing software that was the feature of the corpora. To find out SP of each lexical item, the participants examined the concordance lines with keywords in center. Although, the students frequently used dictionaries to understand the meaning of new vocabulary items, they were asked to direct their attention to collocates of the keyword. Therefore, they did not have to translate every word. After checking the students’ decision, when necessary, the teacher performed a complementary analysis. At this stage, the students compared their own analyses with the teacher’s one and gained knowledge of how to find out the SP of the lexical items appropriately. The teacher gave them feedback by providing commentaries in the class. In the course of time, the students became able to investigate prosodic behavior of items by their own.

As an illustration, when the word TRIGGER was selected from the course book for identifying its SP, its frequent near-synonyms (i.e. BRING ABOUT, CAUSE, PROVOKE, SET OFF, & SPARK) were also investigated. In this synonym set, the words BRING ABOUT and CAUSE had already been investigated by other researchers: the former has a -SP and the latter has a +SP. Therefore, the students were to investigate SP of other members of the synonym set.
Figure 2 shows all 22 selections of instances of the verb TRIGGER in Brown + BNC Written Corpora. The words indicating what entity is ‘triggered’ in each line are highlighted in bold. In most cases, this entity is considered undesirable: fall, breakout, price crisis, revival, trouble, and so on. It can be argued, therefore, that TRIGGER has a -SP.

Figure 2. All 22 selections of instances of the verb TRIGGER in Brown + BNC Written Corpora

All 26 selections of instances of the verb PROVOKE in Brown + BNC Written Corpora are illustrated in Figure 3. In fact, 20 out of the 26 instances of PROVOKE in the corpora have indications of negative attitude. It can be said, then, that PROVOKE has a -SP.
Figure 3. All 26 selections of instances of the verb PROVOKE in Brown + BNC Written Corpora

Negative evaluation of SET OFF is also evident in the examples shown in Figure 4. It shows all 6 selections of instances of the verb SET OFF in Brown + BNC Written Corpora. In most instances, the verb SET OFF co-occurs with items that are evaluatively negative (line 4 in an exception). Therefore, it can be concluded that the word SET OFF has a -SP.

Figure 4. All 6 selections of instances of the verb SET OFF in Brown + BNC Written Corpora

Figure 5 shows all 18 selections of instances of the verb SPARK in Brown + BNC Written Corpora. Form examining its collocations, it can be concluded that the word SPARK has a ~SP.
It is worth noting that SP of near-synonyms was supposed to be examined in a general English environment. Words that participants came across in the course book as well as their near-synonyms were selected as the lexical items whose SP were investigated. In fact, the teacher selected words.

In the present study, the prosodic polarity of a lexical item was determined by comparing the number of instances conveying a positive attitudinal meaning with the number of instances conveying a negative attitudinal meaning. If the positive prosody instances far outnumbered the negative ones, the word was regarded as having a +SP, and vice versa. If there was no tendency towards either prosody, the word was regarded as having a ~SP.

Each session the group worked on two or three words along with their near-synonyms. It should be mentioned that, in order to control the register, near-synonyms from the same register were chosen.

In experimental group B, the participants were informed of the concept of SP. They were asked to find a lexical item’s synonyms in the same thesaurus. However, unlike experimental group A, the teacher, himself, provided them with the SP of lexical items. In other words, they did not undergo the DDL approach.
In the control group, the students were asked to find a lexical item’s near-synonyms in the same thesaurus, too. However, near-synonyms were taught traditionally and without teaching SP of them. Moreover, it was made sure that they had not learnt SP before. In fact, to distinguish between near-synonyms, the participants only used the dictionary definitions and thesaurus.

In order to measure lexical cohesion, before the treatment and immediately after the treatment, IELTS writing tests (task 2) adopted from samples of IELTS writing tests were given to the groups, as pretests and posttests. In both pretest and posttest, the groups were given a time limit of 40 minutes to complete the task. The lexical cohesion of their writing composition was an indicator of their lexical cohesion ability. The participants’ writings were given a score range between 0-9 for their lexical cohesion.

**Data Analysis**

In this study, to compare means of each test within and between the groups, one-way ANOVA and a Post-hoc Tukey Test were used. The null hypotheses of no difference within and between group means were formulated. The alpha level was set to .05.

To decrease subjectivity, two raters assessed the participants’ compositions with respect to lexical cohesion. The average of the scores provided by two raters was considered the participants’ true score.

In order to measure inter-rater reliability for the scores provided by the two raters, the Pearson correlation was used. Table 2 presents the measure of inter-rater reliability of two raters for the pretest scores in all three groups. The correlation achieved in each group was significant. That is, the scores provided by two raters had an acceptable correlation.

<table>
<thead>
<tr>
<th>Groups</th>
<th>N</th>
<th>Pearson Correlation</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>18</td>
<td>.950</td>
<td>.000</td>
</tr>
<tr>
<td>Experimental A</td>
<td>17</td>
<td>.948</td>
<td>.000</td>
</tr>
<tr>
<td>Experimental B</td>
<td>17</td>
<td>.940</td>
<td>.000</td>
</tr>
</tbody>
</table>

Table 3 presents the measure of inter-rater reliability of two raters for the posttest scores in all three groups. The correlation achieved in each group was significant. That is, the scores provided by two raters had acceptable correlations.
Table 3
Inter-rater correlation for the posttest scores

<table>
<thead>
<tr>
<th>Groups</th>
<th>N</th>
<th>Pearson Correlation</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>18</td>
<td>.936</td>
<td>.000</td>
</tr>
<tr>
<td>Experimental A</td>
<td>17</td>
<td>.960</td>
<td>.000</td>
</tr>
<tr>
<td>Experimental B</td>
<td>17</td>
<td>.940</td>
<td>.000</td>
</tr>
</tbody>
</table>

Results

Table 4 shows the descriptive statistics of the participants’ performance in different groups on the pretest.

Table 4
Descriptive statistics (Pretest)

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error</th>
<th>95% Confidence Interval for Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower Bound</td>
</tr>
<tr>
<td>Control</td>
<td>18</td>
<td>5.000</td>
<td>.5145</td>
<td>.1213</td>
<td>4.744</td>
</tr>
<tr>
<td>Experimental A</td>
<td>17</td>
<td>5.206</td>
<td>.5018</td>
<td>.1217</td>
<td>4.948</td>
</tr>
<tr>
<td>Experimental B</td>
<td>17</td>
<td>5.118</td>
<td>.6257</td>
<td>.1518</td>
<td>4.796</td>
</tr>
<tr>
<td>Total</td>
<td>52</td>
<td>5.106</td>
<td>.5454</td>
<td>.0756</td>
<td>4.954</td>
</tr>
</tbody>
</table>

Table 4 indicated that mean and standard deviation of all three groups were to some extent the same. That is, the participants of the three groups performed similarly on the pretest.

In order to find out the difference among the groups on the pretest, One-way ANOVA was performed on the test scores of the three groups’ writing essays. Table 5 shows the results.

Table 5
One-way ANOVA (Pretest)

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>Df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>.374</td>
<td>2</td>
<td>.187</td>
<td>.620</td>
<td>.542</td>
</tr>
<tr>
<td>Within Groups</td>
<td>14.794</td>
<td>49</td>
<td>.302</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>15.168</td>
<td>51</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Based on Table 5 given above, since observed F (0.620) is less than critical F (3.19) with df = 2/49, the difference between the groups is not significant at (p<0.05). Moreover, based on obtained significance of 0.542 which was greater than 0.05, it was concluded that there was no difference between the groups on the pretest. Therefore, all the groups were homogeneous with respect to lexical cohesion at the beginning of the study.

The mean of each group on the pretest is illustrated through a bar graph in Figure 6. It confirms the participants in each group performed almost the same.
Figure 6. The comparison of each group’s mean in cohesion on the pretest

The next step in analyzing the results of the study was the calculation of the students’ scores after the treatment on the posttest. Like the pretest, descriptive and inferential statistics were used for this purpose. The descriptive statistics of participants’ scores on the posttest are given in Table 6.

Table 6
Descriptive statistics (Posttest)

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error</th>
<th>95% Confidence Interval for Mean</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>18</td>
<td>5.139</td>
<td>.6137</td>
<td>.1447</td>
<td>4.834 – 5.444</td>
<td>4.0</td>
<td>6.0</td>
</tr>
<tr>
<td>Experimental A</td>
<td>17</td>
<td>5.882</td>
<td>.4851</td>
<td>.1176</td>
<td>5.633 – 6.132</td>
<td>5.0</td>
<td>6.5</td>
</tr>
<tr>
<td>Experimental B</td>
<td>17</td>
<td>6.000</td>
<td>.4330</td>
<td>.1050</td>
<td>5.777 – 6.223</td>
<td>5.5</td>
<td>7.0</td>
</tr>
<tr>
<td>Total</td>
<td>52</td>
<td>5.663</td>
<td>.6396</td>
<td>.0887</td>
<td>5.485 – 5.842</td>
<td>4.0</td>
<td>7.0</td>
</tr>
</tbody>
</table>

By comparing the mean and standard deviation of each group in Table 6 with the ones of the pretest in Table 4, differences among the groups’ means were arisen. As it is shown in Table 6, the participants’ performances in all groups were changed into greater scores. In order to find out whether the differences in the results of the posttest are significant, one-way ANOVA was used. Table 7 shows the results of this calculation.
Table 7
One-way ANOVA (Posttest)

<table>
<thead>
<tr>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>7.693</td>
<td>2</td>
<td>3.847</td>
<td>14.314</td>
</tr>
<tr>
<td>Within Groups</td>
<td>13.167</td>
<td>49</td>
<td>.269</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>20.861</td>
<td>51</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

According to Table 7, since observed F (14.314) is much greater than critical F (3.19) with df = 2/49, the difference between the groups is significant at (p<0.05). The obtained significance (0.000) is less than significance level set for the study (0.05). Therefore, there is a significant difference between groups and they are not homogeneous on the posttest. Based on descriptive and inferential statistics, the participants performed better on the posttest. In order to find out where the differences exactly exist and compare the groups with each other a Post-hoc Tukey Test was used. The results are presented in Table 8.

Table 8
Post-hoc Tukey Test of multiple comparisons

<table>
<thead>
<tr>
<th>(I) Group</th>
<th>(J) Group</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>Experimental A</td>
<td>-.7435*</td>
<td>.1753</td>
<td>.000</td>
<td>-1.167 to -0.320</td>
</tr>
<tr>
<td></td>
<td>Experimental B</td>
<td>-.8611*</td>
<td>.1753</td>
<td>.000</td>
<td>-1.285 to -0.437</td>
</tr>
<tr>
<td>Experimental A</td>
<td>Control</td>
<td>.7435*</td>
<td>.1753</td>
<td>.000</td>
<td>0.320 to 1.167</td>
</tr>
<tr>
<td></td>
<td>Experimental B</td>
<td>-.1176</td>
<td>.1778</td>
<td>.787</td>
<td>-.547 to 0.312</td>
</tr>
<tr>
<td>Experimental B</td>
<td>Control</td>
<td>.8611*</td>
<td>.1753</td>
<td>.000</td>
<td>0.437 to 1.285</td>
</tr>
<tr>
<td></td>
<td>Experimental A</td>
<td>.1176</td>
<td>.1778</td>
<td>.787</td>
<td>-.312 to 0.547</td>
</tr>
</tbody>
</table>

* The mean difference is significant at the 0.05 level.

Table 8 shows that the groups performed differently in gaining lexical cohesion. The multiple comparisons of the results showed that there was not a significant difference between experimental group A and experimental group B, but there was a significant difference between experimental groups (A & B) and the control group. Thus, it is shown that the participants in the experimental groups performed better on the posttests.

The mean of each group on the posttest is illustrated through a bar graph in Figure 7. It confirms participants in the experimental groups outperformed the control group.
Discussion

From the related literature, it was observed that SP is the main distinguishing feature among near-synonyms. Therefore, lexicographers are recommended to provide necessary information about SP of the lexicon in monolingual dictionaries and apply this knowledge in selecting exact equivalents between the lexicon of different languages in bilingual dictionaries. That is, SP should become an integral part of dictionary definitions to distinguish near-synonyms. It was also found that SP of lexical items is only discovered through the DDL approach.

From investigating the prosodic behavior of synonym sets, two points were discovered: firstly, it was found that in a synonym set, it was not very common to have members with all three types of SP (positive, negative, and neutral). That is, in many cases two types of prosodic behaviors existed in a synonym set. The second point worth mentioning is that negative prosodic behaviors were more frequent than the other two types.

The present study investigated the role of SP in EFL learners’ lexical cohesion. The participants’ initial performance (on the pretest) on lexical cohesion was poor: the mean scores for the control group, experimental group A, and experimental group B were 5.00, 5.206, and 5.118 out of 9, respectively. SP was not taught and practiced in these classes. This
implies that lack of knowledge of SP in traditional classrooms may be one of the causes of not writing coherently enough.

However, after introducing SP and practicing it for a 6-month period, the groups were not homogeneous anymore. On the posttest, the difference between the experimental groups and the control group was significant. In other words, the participants in the experimental groups performed significantly better on the posttest. The results suggested that SP was beneficial for improving lexical cohesion.

Based on the obtained results, the first two null hypotheses were rejected. That is, both teaching SP of lexical items through the DDL approach and making students aware of the SP of lexical items by the teacher have a positive significant effect on EFL learners’ lexical cohesion. However, the last hypothesis was not rejected. That is, there was no significant difference between these two methods of presenting SP of lexical items. Therefore, it can be concluded that SP itself is the main reason of improvement in the learners’ lexical cohesion. In other words, it does not differ if SP of words is taught through DDL or it is provided to the learners by the teacher.

It is quite clear that, concerning discovering SP of lexical items, DDL is more reliable and sophisticated than merely asking a teacher. Therefore, that the DDL approach did not make any difference could be due to the fact that the DDL approach is not convenient, at least in the first experience, it is time-consuming, and it imposes heavy burden on the learner. To overcome this, an attempt should be made to make the DDL approach as friendly as possible.

The most important and basic finding of the present study, emerged from experimental group A, was that it is possible to teach SP of lexical items. That is, the teacher does not have to provide learners with SP of words. The students themselves are able to find out SP of words though the DDL approach.

Another point that is worth mentioning is that as the study took 6 months, it might be concluded that the effect of SP on lexical cohesion is achieved in long term. This is because to achieve considerable difference in lexical cohesion through knowledge of SP, the SP of quite a lot of words should be investigated.

References


